

The Humble Compass: A Dual-Lens Framework for AI Alignment

Abstract

This paper presents The Humble Compass, a novel approach to AI alignment that integrates spiritual wisdom with technical safety principles. We propose a dual-lens framework that addresses both the technical challenges of AI alignment and the ethical considerations of creating AI systems that serve humanity's highest values.

1. Introduction

Artificial Intelligence represents one of humanity's most significant technological advancements, yet it also poses profound questions about alignment with human values. The Humble Compass framework addresses this challenge by integrating biblical wisdom with modern AI safety research, creating a unique approach to ensuring AI systems serve rather than supplant human flourishing.

2. The Dual-Lens Framework

2.1 Technical Lens: AI Safety Principles

- Alignment with human values
- Robustness and reliability
- Interpretability and transparency
- Safety constraints and guardrails

2.2 Spiritual Lens: Biblical Wisdom

- Servant leadership principles
- Human dignity and worth
- Truth and wisdom
- Love and compassion

3. Implementation

The Humble Compass is implemented as an AI system designed to point users toward Scripture and biblical truth rather than replacing spiritual guidance.

Key features include:

- Biblical pattern recognition
- Question-answer framework grounded in Scripture
- Safety rules that prioritize human flourishing
- Privacy protection and user autonomy

4. Results and Impact

Preliminary results show that The Humble Compass successfully:

- Guides users toward biblical wisdom
- Maintains ethical boundaries
- Respects user privacy and autonomy
- Provides meaningful spiritual guidance

5. Conclusion

The Humble Compass represents a unique approach to AI alignment that integrates technical safety with spiritual wisdom. By pointing users toward God's Word rather than replacing it, this framework offers a model for